

基于文本挖掘的临床进修反馈分析与改进探索

刘 蕾¹ 张 越¹ 王 蕾² 蒋 俭¹

(¹首都医科大学附属北京同仁医院 北京 100730

²中国医学科学院/北京协和医学院医学信息研究所 北京 100020)

〔摘要〕 **目的/意义** 评估临床进修现状，为精准改进教学管理提供实证依据。**方法/过程** 收集进修医师评价文本，采用BERTopic与中文预训练语言模型开展主题与情感分析，结合改进的重要性-绩效分析方法构建“主题关注度-情感满意度”矩阵，对低满意度主题的负向文本进行关键词共现分析。**结果/结论** 共识别出6个主题，分别划分至优势巩固、关键改进、适度提升、次要优化4个象限，建议在进修教学管理中巩固师资优势、增加临床与科研实践机会、优化管理流程与学习支持环境。

〔关键词〕 临床进修；文本挖掘；教学管理

〔中图分类号〕 R-058 **〔文献标识码〕** A **〔DOI〕** 10.3969/j.issn.1673-6036.2026.02.009

Analysis and Improvement Strategies of Clinical Fellowship Feedback Based on Text Mining

LIU Lei¹, ZHANG Yue¹, WANG Lei², JIANG Jian¹

¹Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China; ²Institute of Medical Information, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100020, China

〔Abstract〕 **Purpose/Significance** To assess the current status of clinical fellowship, and to provide empirical evidence for targeted improvement of teaching management. **Method/Process** Textual evaluation data are collected from clinical fellows. BERTopic and a Chinese pre-trained language model are used to perform topic modeling and sentiment analysis. An improved importance-performance analysis approach is then applied to construct a “topic attention-sentiment satisfaction” matrix, and keyword co-occurrence analysis is conducted on negative texts of low-satisfaction topics. **Result/Conclusion** Six topics are identified and classified into four quadrants: strengths to be maintained, key areas for improvement, areas for moderate enhancement, and secondary optimization areas. It is recommended that clinical fellowship management focus on consolidating faculty strengths, expanding clinical and research opportunities, and optimizing both management processes and the learning support environment.

〔Keywords〕 clinical fellowship; text mining; teaching management

〔修回日期〕 2025-11-29

〔作者简介〕 刘蕾，助理研究员，发表论文5篇；通信作者：蒋俭，主任医师，副教授。

〔基金项目〕 首都医科大学附属北京同仁医院青年管理人才项目（项目编号：2023-YJJ-ZZL-064）。

1 引言

临床进修教育是提升医师临床实践能力的关键环节，科学评估其培训效果至关重要^[1]。既有研究^[2-4]多围绕进修项目的结构设置、满意度与教学质量等开展问卷调查，在推进评价指标体系建设和量化比较方面发挥了重要作用。然而，此类研究多依赖预设量表与固定题项，侧重总体满意度或有限维度的主观评分，难以及时、充分地捕捉进修医师在真实情境中的学习经历、个体诉求及潜在问题。与此同时，教育领域已有研究^[5]引入文本挖掘与自然语言处理技术，对课程评价、教学反馈及在线学习评论等非结构化文本进行分析，证实文本挖掘能够在学习者情感识别、潜在问题预警和改进建议提炼等方面弥补传统方法的不足。但针对临床医师进修反馈评价开展的相关研究相对匮乏，鲜有从整体视角全面揭示进修体验的构成维度及情感倾向的实证工作。本研究聚焦进修医师评价文本，综合运用主题模型与情感分析方法，系统挖掘进修医师的关注重点与满意度情况，识别现阶段优势与薄弱环节，为精准优化临床进修教学与管理提供数据支持与决策参考。

2 资料与方法

2.1 数据来源与预处理

收集首都医科大学附属北京同仁医院 2023 年 9 月 28 日—2025 年 9 月 29 日结业进修医师填写的“进修工作情况调查表”共 872 份，提取其中“您对本次进修经历的评价”开放式文本题项回答数据，共获得有效文本 869 条。所有数据均在匿名条件下收集，研究仅用于教学管理改进，不涉及个人隐私信息。使用 Python 3.12 对原始自然语言文本进行数据清洗，删除无意义字符、表情符号、重复文本后，最终纳入 861 条文本数据开展研究分析。

2.2 主题识别

采用 BERTopic 模型进行评价文本主题识别^[6]。选

用 paraphrase-multilingual-MiniLM-L12-v2 作为预训练模型，将预处理后的文本进行语义向量化；利用均匀流形近似与投影（uniform manifold approximation and projection, UMAP）算法对向量进行降维，利用层次密度聚类（hierarchical density-based spatial clustering of applications with noise, HDBSCAN）算法对语义相似文本进行聚类；利用基于类别的词频-逆文档频率（class-based term frequency-inverse document frequency, c-TF-IDF）计算各簇的关键词权重，以提取代表性术语。在此基础上，结合高频词与随机抽样的原始文本，对聚类结果进行人工审阅与语义标注，合并语义重叠主题，最终确定核心主题名称和分类结果。

2.3 情感分析

基于 PyTorch 框架与 Hugging Face Transformers 库，加载中文预训练语言模型 hfl/chinese-roberta-wwm-ext-large 进行情感分类，模型输出类别包括正向、中性与负向。从全部评价文本中随机抽取 150 条样本进行人工情感标注^[7]。考虑到标注样本量有限且不同情感类别分布不均，为降低单次训练、测试划分带来的偶然性，并保证各类别在验证集中的代表性，对标注文本采用 5 折分层交叉验证，在每一折中以 4/5 的样本作为训练集、1/5 的样本作为验证集。在模型微调阶段，引入 Focal Loss 损失函数缓解类别不平衡问题，并基于各折验证集上的表现进行超参数调优。采用准确率（accuracy）和宏平均 $F1$ （macro- $F1$ ）作为评价指标，计算 5 折结果的平均值及标准差，以衡量模型在小样本情感识别任务中的性能与稳定性。基于交叉验证中全部验证集的预测结果，构建 3 类情感的混淆矩阵，用于分析模型在不同情感类别上的识别差异。

2.4 重要性-绩效分析

借鉴改进的重要性-绩效分析（importance-performance analysis, IPA）框架^[8]，构建“主题关注度-情感满意度”矩阵，确定改进优先级。基于 BERTopic 模型输出的主题概率分布计算主题关注度 θ_k 。其中， D 为文本总数， K 为主题总数， $P_{d,k}$ 表示第 d 条文本属于主题 k 的概率。通过聚合各主题

下的情感分类结果计算情感满意度 α_k 。其中， $P_k^{(正向)}$ 为主题 k 下正向情绪文本数， $P_k^{(总)}$ 为主题 k 下的总文本数。

$$\theta_k = \frac{\sum_{d=1}^D P_{d,k}}{\sum_{k=1}^K \sum_{d=1}^D P_{d,k}} \quad (1)$$

$$\alpha_k = \frac{P_k^{(正向)}}{P_k^{(总)}} \quad (2)$$

鉴于本研究中关注度与情感满意度均为 0~1 区间的比例型指标，且存在偏态分布和极端值，中位数相较均值具有更好的稳健性^[9]，采用所有主题关注度和情感满意度的中位数作为划分阈值，将主题映射至“高关注-高满意”“高关注-低满意”“低关注-高满意”“低关注-低满意”4个象限，以刻画各主题的相对位置并界定管理改进优先级。

2.5 关键词共现网络分析

为揭示低满意度主题涉及的问题，对其中负向情绪文本开展关键词共现网络分析。针对情绪标注为负向的文本，构建自定义领域词典，利用jieba分词并去除停用词，以提高分词准确性。统计任意两个关键词在同一文本中的共现次数，构建关键词共现矩阵与加权网络，并设置共现次数阈值为3，仅保留不低于该阈值的关键词共现边，以减少噪声。采用社区检测算法 Louvain 对网络进行聚类，识别关键词高频组合，提炼待改进的“问题簇”。

3 结果

3.1 进修医师评价主题与情感倾向

BERTopic 主题建模结果提示，有 30 条 (3.48%) 评价文本属离群值类、145 条 (16.84%) 文本属“无评价”类、95 条 (11.03%) 文本属“满意或感谢”短文本类、105 条 (12.20%) 文本属“祝福或期待”等语义类，除此之外的 486 条 (56.45%) 文本共涉及 6 类主题：“信息设备”“教师评价”“收获体验”“教学培训”“后勤保障”和“进修管理”。

情感分析模型的 5 折平均准确率为 0.9067±0.0327，macro-F1 为 0.9045±0.0395，说明在小样本条件下，模型在不同交叉验证折次之间的性能波动较小，具有

一定稳定性和鲁棒性。基于交叉验证中全部验证集的预测结果，构建正向、中性和负向 3 类情感的混淆矩阵，见表 1。模型对 3 类情感均具有较好的识别能力，仅有少量样本类别边界模糊。该模型共识别出正向文本 236 条 (48.56%)、中性文本 31 条 (6.38%)、负向文本 219 条 (45.06%)，见表 2。

表 1 情感分类混淆矩阵(条)

类别	预测_正向	预测_中性	预测_负向
真实_正向	80	1	2
真实_中性	1	15	1
真实_负向	7	1	42

表 2 进修评价文本主题类别与情感分类

主题名称	文本数(条)	占比(%)	特征词	情感类型	情感文本数量(条)	相对占比(%)
信息设备	35	4.07	医嘱、病历、系统、改善、无纸化	正向	0	0
				中性	0	0
				负向	35	100
教师评价	162	18.82	老师、主任、带教、认真负责、指导	正向	145	89.51
				中性	10	6.17
				负向	7	4.32
收获体验	105	12.20	收获、提高、机会、专业知识、满意	正向	85	80.95
				中性	15	14.29
				负向	5	4.76
教学培训	103	11.96	理论知识、临床实践、科研、讲座、手术	正向	2	1.94
				中性	4	3.88
				负向	97	94.17
后勤保障	21	2.44	住宿、食堂、报销、劳务费、方便	正向	0	0
				中性	0	0
				负向	21	100
进修管理	60	6.97	名额、灵活、流程、态度、烦琐	正向	4	6.67
				中性	2	3.33
				负向	54	90.00

3.2 主题关注度与情感满意度分析结果

6 类主题的主题关注度范围为 0.04~0.35，情感满意度范围为 0~0.90。以所有主题关注度的中位数 0.15 和情感满意度的中位数 0.04 作为分割阈值，将主题划分并命名为“优势巩固”“关键改进”“适度提升”和“次要优化”4 个象限，见图 1。“教师评价”($\theta=0.35, \alpha=0.90$)与“收获体验”($\theta=0.23, \alpha=0.81$)位于优势巩固区，提示教师水平和整体进修收获均具有较高关注度和较好满意度；“教学培训”

握制度规范、诊疗流程、应急处理、信息系统应用等内容,从而缩短适应周期^[15]。三是科研能力建设方面,当前科研培训与临床轮转之间存在脱节现象。应在进修周期内针对科研设计、论文撰写等薄弱环节,设置契合进修医师特点的科研培训系列课程,提出可实现的阶段性目标与连续性支持机制,如围绕既有临床病例或科室在研项目,引导进修医师完成文献综述、研究构思或病例报告等不同难度的科研任务,并通过集中授课或小组讨论形式,提供过程性反馈,增加进修医师接触科研的机会^[16]。四是教学模式创新方面,虽然多数医院已开展教学查房和病例讨论,但文本反馈提示其临床思维训练的实际效果有限。建议丰富结构化教学活动比重,围绕典型病例设置标准化诊疗决策节点,引导进修医师阐述思路,再由高年资医师进行聚焦式点评。与简单增加教学活动数量相比,这种基于问题与决策过程的结构化调整,更有助于提升满意度与教学效果^[17]。

4.3 管理流程与效率有待提升

“进修管理”主题反映了行政流程效率、岗位工作安排和信息沟通响应3方面问题,提示以行政管理为中心的模式难以充分回应进修医师的体验需求。在有限资源约束下,管理优化的关键在于通过流程再造与信息化手段,提升管理效率与学习者参与体验。在行政流程效率方面,评价文本集中反映了进修管理路径烦琐的问题。应升级进修医师全流程管理系统^[18],通过流程可视化和节点前移提升感知效率。将涉及的主要管理节点以流程图或时间轴形式呈现,使进修医师在学习过程中全面了解每个关键环节及所需材料,从而减少反复咨询与线下办理^[18]。在工作安排与时间分配方面,“临床任务重、学习时间少”的反馈比较突出,揭示了进修医师角色界定模糊的现状。应深入探讨进修医师嵌入式岗位工作模式^[19],按比例分配临床服务与参训时间,帮助其改善时间分配失衡状况,以便有余力从事学术或科研活动。另外,通过在排班和考核中适度强化“学习任务”,推动临床科室从“以人手补充为主”转向“以培养目标为导向”的岗位使用逻辑。

在信息沟通响应方面,进修医师接收各类信息处于相对被动状态,应将进修管理由单向告知转为双向、持续的互动机制。一方面,通过统一在线平台整合培训安排、值班要求、考核信息等关键通知,减少依赖工作群等方式分散发布消息;另一方面,建立固定渠道收集进修医师在轮转安排、学习资源等方面的意见,并对共性问题进行定期反馈和说明,从而增强管理的透明度和应答性。

4.4 支持环境有待协同优化

“后勤保障”与“信息设备”两个主题虽然对整体进修体验的影响有限,但作为培训环境建设的支撑环节,可能通过累积效应削弱教学与管理改进的成效。在工作环境与生活保障方面,部分进修医师提及提供住宿、改善环境、提升就餐体验等需求,据此应在有限资源内适度提升环境的舒适度与后勤保障的可及性。可推行食宿安排与生活设施信息透明化,使进修医师对即将面对的生活环境有合理预期,减轻落差感。提高进修医师生活补助,加强人文关怀,以减轻其经济与情绪压力。在信息化支持方面,医院信息系统及信息设备有较大提升空间。同时,应提高现有工作和学习平台对进修医师的开放程度,重点提升文献以及线上教学等资源的可及性,并优化访问流程与使用体验。

5 结语

基于文本挖掘的临床进修反馈分析,有效识别出进修培训的优势与改进重点,验证了该方法在教学评估中的应用价值。本研究尚存在一定局限性。一是情感分析方法依赖预训练模型,且当前人工标注样本量较少,可能导致分析偏差。后续研究可通过扩大标注规模或引入主动学习策略进一步提升模型性能。二是研究数据仅来源于一家三甲教学医院的横断面评价文本,结果主要反映了该院进修教学管理的现状特征,结果的外部代表性有待进一步验证。未来可通过多渠道数据采集和多维度研究方法,完善分析框架并开展针对性干预研究,以验证改进措施的效果,推动临床进修培训持续优化。

作者贡献：刘蕾、张越、王蕾负责研究设计、数据分析、论文撰写；蒋俭负责提供指导、论文修订。
利益声明：所有作者均声明不存在利益冲突。

参考文献

- 张潇潇, 李娟, 刘扬, 等. 某医院血液科进修医师培训满意度影响因素研究[J]. 中华医学教育探索杂志, 2020, 19(11): 1328-1330.
- 王婷, 马晓, 金铭, 等. 三级综合医院进修医师满意度及需求调查[J]. 中国公共卫生, 2017, 32(10): 1515-1518.
- 张潇潇, 曾庆奇, 侯丽如, 等. 进修医师培训质量满意度影响因素研究[J]. 中国卫生经济, 2022, 41(12): 98-100, 109.
- 蒙世佼, 霍墨菲, 张家亮. 脑血管病专科进修医师工作满意度与培训成效分析及改进策略研究[J]. 中国医刊, 2025, 60(5): 613-616.
- 林铭炜, 许江松, 林佳胤, 等. 面向在线教育的学习者情感识别综述[J]. 控制与决策, 2024, 39(4): 1057-1074.
- EGGER R, YU J. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify twitter posts[J]. Frontiers in sociology, 2022, 7(5): 886498.
- GAO T, FISCH A, CHEN D. Making pre-trained language models better few-shot learners[C]. Online: The 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, 2021.
- 毕达天, 王璐, 王雨菲, 等. AIGC 产品用户需求特征体系构建及改进策略研究[J]. 图书情报工作, 2024, 68(14): 14-24.
- BACON D R. A comparison of approaches to importance-performance analysis [J]. International journal of market research, 2003, 45(1): 1-15.
- 宋泽亮, 孙嫣然, 杨红军, 等. 基于一院多区管理模式下进修医师培训现状与管理对策分析[J]. 中国医药导报, 2025, 22(3): 107-110.
- 王燕, 袁佳艳, 张菲菲. 我国教学医院进修医生管理存在的问题及对策[J]. 医学与社会, 2019, 32(1): 127-129.
- FU Y, ZHAO G, YU C, et al. Study on strategies to improve clinical fellow training satisfaction based on importance-performance analysis (IPA) [J]. BMC medical education, 2025, 25(1): 1137.
- 马满姣, 马璐璐, 王锐颖, 等. 麻醉科进修医师临床教学培训模式的探索研究[J]. 协和医学杂志, 2024, 15(2): 462-465.
- 罗林枝, 李琦, 罗静, 等. 北京协和医院进修医师院前培训在线学习模式探索与学习效果评估[J]. 基础医学与临床, 2017, 37(3): 431-434.
- 张佩, 胡亚洲. 肿瘤专科医院进修培训模式及改进思考[J]. 医院管理论坛, 2021, 38(8): 77-79, 74.
- 马璐璐, 王锐颖, 张志永, 等. 北京协和医院麻醉科进修医师科研规范化培训模式探索[J]. 中国医药导报, 2024, 21(10): 50-53.
- FRATUS A, NIRUNSUK P, FENG X, et al. A survey on the perspective and benefits of being an international visiting surgeon at a leading level I trauma centre [J]. International orthopaedics, 2025, 49(1): 29-33.
- 李欢, 周磊. 三级医院进修生全流程管理系统的设计与实现[J]. 中国数字医学, 2022, 17(5): 35-39.
- GALLOWAY R, CASTLE J, BROWN A, et al. An evaluation of clinical fellow programmes in an acute teaching hospital trust [J]. British journal of hospital medicine, 2023, 84(10): 1-6.

2026 年《医学信息学杂志》征订启事

《医学信息学杂志》是国内医学信息领域创刊最早的国家级期刊。主管单位：国家卫生健康委员会；主办单位：中国医学科学院。中国科技核心期刊（中国科技论文统计源期刊），美国《化学文摘》、《乌利希期刊指南》及 WHO 西太区医学索引（WPRIM）收录，并收录于国内 3 大数据库。主要栏目：专论，医学信息研究，医学信息技术，医学信息资源管理与利用，医学信息教育等。读者对象：医学信息领域专家学者、管理者、实践者，高等院校相关专业的师生及广大医教研人员。

《医学信息学杂志》国内外公开发行，每册定价：20 元（月刊），全年 240 元。邮发代号：2-664，全国各地邮局均可订阅。也可到编辑部订购：北京市朝阳区雅宝路 3 号中国医学科学院医学信息研究所《医学信息学杂志》编辑部（100020）；电话：010-52328686, 52328687。

《医学信息学杂志》编辑部